REMARKS

Claims 1-16, and 19-22 are pending. The Examiner's reconsideration of the objections and rejections is respectfully requested in view of the remarks.

Claims 1-10 and 14-22 have been rejected under 35 U.S.C. 103(a) as being unpatentable over DeCosta et al. (USPN 6,826,553) in view of Gardner et al. (US Pub. Patent 2003/0188112). The Examiner stated essentially that the combined teachings of DeCosta and Gardner teach or suggest all the limitations of claims 1-10 and 14-22.

Claims 1, 19, and 20 are the independent claims.

Claims 1 and 19 claim, *inter alia*, "identifying a first set of attribute occurrences in the template generated semi-structured document using an ontology; determining a boundary of each multi-attribute data record in the template generated semi-structured document; learning a pattern for an attribute corresponding to an identified attribute occurrence of the first set in the template generated semi-structured document; and applying the pattern within the boundary of each multi-attribute data record in the template generated semi-structured document to extract a second set of attribute occurrences." Claim 20 claims, *inter alia*, "an ontology for identifying a first set of attribute occurrences in the template generated semi-structured

10

document, the ontology comprising a set of concepts and a set of attributes associated with every concept; a boundary module for determining a boundary of each multi-attribute data record in the template generated semi-structured document; and a pattern module for learning a pattern for an attribute corresponding to an identified attribute occurrence of the first set in the template generated semi-structured document."

DeCosta teaches a system for automatically extracting data from at least one electronic document (see paragraph Abstract). DeCosta does not teach or suggest "identifying a first set of attribute occurrences in the template generated semi-structured document using an ontology" "and applying the pattern within the boundary of each multi-attribute data record in the template generated semi-structured document to extract a second set of attribute occurrences" as claimed in Claims 1 and 19, nor "an ontology for identifying a first set of attribute occurrences" "and a pattern module for learning a pattern for an attribute corresponding to an identified attribute occurrence of the first set in the template generated semi-structured document" as claimed in Claim 20. DeCosta's method navigates to a web page according to a recording of previous user navigation (via a navigation recording module 12) (see col. 5, lines 37-51), and implements pattern matching to retrieve relevant information

from a navigated to web site based on user selections (via an extraction recording module 22) (see col. 5, lines 55-67). DeCosta's pattern matching operates as the sole means of extracting information. Further, the pattern matching is based on a recording of user activity. Nowhere does DeCosta teach or suggest pattern matching to extract attribute occurrences based on attribute occurrences identified using an ontology, essentially as claimed in Claims 1, 19, and 20. The pattern matching of DeCosta is preformed based on user activity. User activity is not an ontology, essentially as claimed in Claims 1, 19, and 20. Therefore, DeCosta fails to teach or suggest all the limitations of Claims 1, 19, and 20.

Gardner teaches that an ontology may be used to enable effective syntactic and semantic mapping between any number of different entities (see paragraph [0051]). Gardner does not teach or suggest "identifying a first set of attribute occurrences in the template generated semi-structured document using an ontology" "and applying the pattern within the boundary of each multi-attribute data record in the template generated semi-structured document to extract a second set of attribute occurrences" as claimed in Claims 1 and 19, nor "an ontology for identifying a first set of attribute occurrences" "and a pattern module for learning a pattern for an attribute corresponding to

an identified attribute occurrence of the first set in the template generated semi-structured document" as claimed in Claim 20. Gardner teaches only an ontology system. Gardner does not teach or suggest pattern matching to extract attribute occurrences based on attribute occurrences identified using an ontology, essentially as claimed in Claims 1, 19, and 20. Therefore, Gardner fails to cure the deficiencies of DeCosta.

The combined teachings of DeCosta and Gardner fail to teach or suggest learning a pattern based on attribute occurrences identified using an ontology. DeCosta teaches learning only recording user activity. Gardner teaches only that an ontology system accesses information form data sources (for example, see paragraph [0062]). Neither DeCosta nor Gardner teach or suggest learning a pattern from information identified using an ontology.

Further still, combined teachings of DeCosta and Gardner fail to teach or suggest applying two extraction techniques to a template generated semi-structured document for extracting a first and second set of attribute occurrences – at most DeCosta and Gardner teach applying only one method for extracting information.

In view of the above, it is respectfully submitted that the combined teachings of DeCosta and Gardner fail to teach or

suggest "identifying a first set of attribute occurrences in the template generated semi-structured document using an ontology" "and applying the pattern within the boundary of each multi-attribute data record in the template generated semi-structured document to extract a second set of attribute occurrences" as claimed in Claims 1 and 19, or "an ontology for identifying a first set of attribute occurrences" "and a pattern module for learning a pattern for an attribute corresponding to an identified attribute occurrence of the first set in the template generated semi-structured document" as claimed in Claim 20.

Claims 2-9 depend from Claim 1. Claims 21 and 22 depend from Claim 20. The dependent claims are believed to be allowable for at least the reasons given for Claims 1, 19, and 20. Claims 17 and 18 have been cancelled. At least Claim 7 is believed to be allowable for additional reasons.

Claim 7 claims, "further comprising the step of generalizing the pattern of the attribute occurrence prior to applying the pattern."

DeCosta does not teach or suggest, "generalizing the pattern of the attribute occurrence prior to applying the pattern" as claimed in Claim 7. DeCosta records patterns of user activity (see col. 9, line 5 to col. 10, lines 17). As more user selections are considered, the patterns of DeCosta become more

predictable (see col. 9, lines 8-12). Predictability is desired of DeCosta's patterns. Thus, DeCosta teaches away from generalizing patterns, as a generalization may make the patterns less predictable. Indeed, nowhere does DeCosta teach or suggest generalizing the pattern of the attribute, essentially as claimed in Claim 7.

Gardner teaches combining multiple available ontologies (see paragraph [0105]). An ontology is not a pattern as claimed in Claim 7. According to Claim 1, the ontology is used to identify a first set of attribute occurrences and a pattern is used to extract a second set of attribute occurrences. The terms ontology and pattern are not analogous. Gardner teaches only combining ontologies. Gardner fails to teach or suggest generalizing the pattern of the attribute, essentially as claim in Claim 7. Therefore, Gardner fails to cure the deficiencies of DeCosta.

The combined teachings of DeCosta and Gardner fail to teach or suggest, "generalizing the pattern of the attribute occurrence prior to applying the pattern" as claimed in Claim 7.

The Examiner's reconsideration of the rejection is respectfully requested.

Claims 11 and 13 have been rejected under 35 U.S.C. 103(a) as being unpatentable over DeCosta in view of Gardner, and

further in view of Oommen et al. (US Patent Pub. 2003/0195890).
The Examiner stated essentially that the combined teachings of
DeCosta, Gardner, and Oommen teach or suggest all the
limitations of Claims 11 and 13.

Claims 11 and 13 depend from Claim 1. The dependent claims
are believed to be allowable for at least the reasons given for
Claim 1. The Examiner's reconsideration of the rejection is
respectfully requested.

Claim 12 has been rejected to under 35 U.S.C. 103(a) as
being unpatentable over DeCosta in view of Gardner, and further
in view of Oommen, and further in view of Bruno (Efficient
Creation of Statistics over Query Expressions, ICDE 2003,
Bangalore, India, March 5-8, 2003). The Examiner stated
essentially that the combined teachings of DeCosta, Gardner,
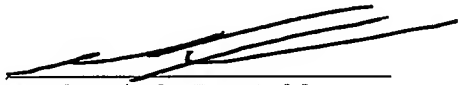Oommen, and Bruno teach or suggest all the limitations of Claim
12.

Claim 12 depends from Claim 1. The dependent claim is
believed to be allowable for at least the reasons given for
Claim 1. The Examiner's reconsideration of the rejection is
respectfully requested.

For the forgoing reasons, the application, including Claims
1-16 and 19-22, is believed to be in condition for allowance.

Early and favorable reconsideration of the case is respectfully requested.

Respectfully submitted,

By:

Nathaniel T. Wallace
Reg. No. 48,909
Attorney for Applicant(s)

Mailing Address:
F. CHAU & ASSOCIATES, LLC
130 Woodbury Road
Woodbury, New York 11797
TEL: (516) 692-8888
FAX: (516) 692-8889